Gauge Invariance, Finite Temperature, and Parity Anomaly in $D = 3$

S. Deser,[1] L. Griguolo,[2] and D. Seminara[1]
[1]Department of Physics, Brandeis University, Waltham, Massachusetts 02254
[2]Center for Theoretical Physics, Laboratory for Nuclear Science and Department of Physics, Massachusetts Institute of Technology, Cambridge, Massachusetts 02139
(Received 8 May 1997)

The effective gauge field actions generated by charged fermions in QED_3 and QCD_3 can be made invariant under both small and large gauge transformations at any temperature by suitable regularization of the Dirac operator determinant, at the price of parity anomalies. We resolve the paradox that the perturbative expansion is not invariant, as manifested by the temperature dependence of the induced Chern-Simons term, by showing that large (unlike small) transformations and hence their Ward identities are not perturbative order preserving. Our results are illustrated through concrete examples of field configurations, where the interplay between gauge and parity anomalies is also exhibited. [S0031-9007(97)04005-2]

PACS numbers: 11.10.Wx, 11.10.Kk, 11.15.Bt, 11.30.Rd

Three-dimensional gauge theories are of physical interest in the condensed matter context [1], but display special features requiring understanding different from their four-dimensional counterparts. In particular, we will be concerned with the complex of problems associated with the presence of Chern-Simons (CS) terms [2], the necessary quantization of their coefficients [2,3] in the action stemming from the possibility of making homotopically nontrivial *"large"* gauge transformations, and the effect of quantum loop corrections on this sector [4–6]. While large transformations are always relevant in the non-Abelian case, they also come into play in the physically most interesting case of QED_3 at finite temperatures where the compactified Euclidean time/temperature provides a nontrivial S^1 geometry. These exotic features have been the subject of a large literature [7], as they seemingly lead to a paradox: on the one hand, large gauge invariance appears to require quantization of the CS term's coefficient; on the other, matter loop contributions to the effective gauge field action at finite temperatures yield a perturbative expansion in which the CS term acquires temperature-dependent, hence nonquantized, coefficients that seem to signal a gauge anomaly. This is particularly puzzling since both the matter action and the process of integrating out its excitations should be intrinsically gauge invariant. We will establish that the effective action is indeed invariant under both small and large transformations using the classic results of [8] that gave a clear definition of the Dirac operator's functional determinant by means of ζ-function regularization. Instead, we will see that it is the perturbative expansion that is noninvariant because large transformations necessarily introduce nonanalytic dependence on the charge so that expansion in e^2 and large gauge invariance are mutually incompatible: the induced Chern-Simons term's noninvariance is precisely compensated by further nonlocal contributions in the effective action. We will also note the necessary clash between gauge invariance and parity conservation, similar to that in the familiar axial anomaly in even dimensions. All these features are illustrated in detail by explicit consideration of some nontrivial configurations that enable us to "parametrize" the Chern-Simons aspects in both the Abelian and non-Abelian context.

Let us begin with the peculiar properties of large gauge transformations that invalidate the usual Ward identity counting. For U(1) in particular, and restoring (for the moment) explicit dependence on e, we have $A_\mu \longrightarrow A_\mu + e^{-1}\partial_\mu f$. Normally, we can merely redefine $\tilde{f} = e^{-1}f$. This is also true at finite temperature for the small gauge transformations since f is required to be periodic only in Euclidean time $\beta = (\kappa T)^{-1}$. Thus a perturbative expansion will be small gauge invariant order by order. But for large ones, the periodicity condition becomes $f(0,\mathbf{r}) = f(\beta,\mathbf{r}) + 2\pi n$, with $n \in \mathbb{Z}$, and a rescaling will merely hide the e^{-1} factor in the boundary conditions, leaving the large shift $A_0 \longrightarrow A_0 + 2\pi n/e$ unaffected. This intrinsic dependence means that only the *full* effective action (as we will show), but not its individual expansion terms (including CS parts) will remain invariant. [Perturbative noninvariance will also characterize any other expansion that fails to commute with the above boundary condition.] We are therefore driven to a careful treatment of the induced effective action $\Gamma[A]$ resulting from integrating out the charged matter, for us massive fermions, according to the usual relation $\exp(-\Gamma[A]) = \det(i\not{D} + im)$ where D_μ is the U(1) covariant derivative. The extension to N flavors and to the non-Abelian case will be seen to be straightforward.

Our three-space has $S^1(\text{time}) \times \Sigma$ topology, Σ being a compact Riemann two-surface such as a sphere S^2 or a torus T^2, depending on the desired spatial boundary conditions. We work with a finite two-volume in order to avoid infrared divergences associated with the continuous spectrum in an open space. Before proceeding, let us see how gauge invariance constrains the form of the determinant. [To avoid irrelevant spatial homotopies, we

 0031-9007/97/79(11)/1976(4)$10.00

shall here take Σ to be the sphere.] Because of the existence of the nontrivial S^1 cycle we can construct (besides $F_{\mu\nu}$) the gauge invariant holonomy $\Omega(\mathbf{r}) \equiv \exp[i\int_0^\beta A_0(t',\mathbf{r})\,dt']$. The new information carried by Ω is encoded entirely in a topological degree of freedom that inherits the nontrivial behavior of A_0 under large gauge transformations. Since Ω is unimodular and obeys $\nabla\Omega = i\Omega\int_0^\beta \mathbf{E}(t',\mathbf{r})\,dt'$, it follows that it is the product of a nonlocal functional of $\mathbf{E}$ and of the two-geometry times a phase $\exp(2\pi i a)$. We may take the constant a, called the flat connection, to represent this new degree of freedom; it transforms according to $a \longrightarrow a + 1$ under large transformations. Consequently, the determinant is a functional of both $F_{\mu\nu}$ and a, obeying the additional Ward identity $e^{-\Gamma(a+1,F_{\mu\nu})} = e^{-\Gamma(a,F_{\mu\nu})}$: it is periodic in a (actually also Γ is). Then we may Fourier expand it

$$\begin{aligned}\exp[-\Gamma(F_{\mu\nu},a)] &= \sum_{k=-\infty}^{\infty}[\hat{\Gamma}_k^{(1)}(F_{\mu\nu})\cos 2\pi ka + \hat{\Gamma}_k^{(2)}(F_{\mu\nu})\sin 2\pi ka] \\ &= e^{iI_{\mathrm{CS}}}\sum_{k=-\infty}^{\infty}\{\Gamma_k^{(1)}(F_{\mu\nu})\cos\pi[2k-\Phi(F)]a + \Gamma_k^{(2)}(F_{\mu\nu})\sin\pi[2k-\Phi(F)]a\}, \end{aligned} \tag{1}$$

where $\Phi(F) = (1/4\pi)\int d^2x\,\epsilon^{ij}F_{ij}$ is the electromagnetic flux through S^2 and $I_{\mathrm{CS}} = (1/4\pi)\int(dx)\epsilon^{\mu\nu\rho}A_\mu\partial_\nu A_\rho$. To write this representation of the effective action we have used the fact that Chern-Simons action I_{CS} can be rewritten as $\pi a\Phi(F)$ plus a functional of F only. [Effectively, we represent the "large" aspects through I_{CS}, or a, and the "small" ones through $F_{\mu\nu}$.] In the second equality, we have factored out an explicit CS part, which is also the intrinsic parity anomaly, as we shall see; the second form will be realized in our explicit examples below. These two representations also make clearer how explicit CS terms can be present without loss of gauge invariance.

We now return to the definition of the effective action. Within our framework, the Dirac operator is a well-defined elliptic operator [8] whose determinant can be rigorously and uniquely specified. The ζ-function regularization [9] defines the formal product of all the eigenvalues λ_n as

$$\det i(\not{D} + m) = \Pi\lambda_n \equiv \exp[-\zeta'(0)], \qquad \zeta(s) \equiv \sum(\lambda_n)^{-s} \tag{2}$$

with implicit repetition over degenerate eigenvalues. For $s > 3$ in $D = 3$ [8], the above series converges and its analytic extension defines a meromorphic function with only simple poles. It is regular at $s = 0$, thereby assuring the meaningfulness of (2). A careful definition of λ_n^{-s} is required to avoid ambiguities. We take it to be $\exp(-s\ln\lambda_n)$ where the cut is chosen to be over the positive real axis, $0 \le \arg\lambda_n < 2\pi$, enabling us to rewrite $\zeta(s)$ in the more convenient form

$$\zeta(s) = \sum_{\mathrm{Re}\,\lambda_n>0}(\lambda_n)^{-s} + \exp(-i\pi s)\sum_{\mathrm{Re}\,\lambda_n<0}(-\lambda_n)^{-s}. \tag{3}$$

Changing the cut alters the determinant only if it intersects the line $\mathrm{Im}\,z = m$, in which case the only relevant difference is the sign of the exponential in (3). This alternative choice does not affect gauge invariance, but does change the sign of the parity anomaly terms in $\Gamma[A]$ as was noted in [10] by more complicated considerations. Once the determinant of the Dirac operator has been regularized, its full gauge invariance reduces to that of its eigenvalue spectrum. But small transformations do not affect the λ_n at all, while the large ones merely permute them, as in usual illustrations of index theorems [11]; every well-defined symmetric function of the spectrum, such as $\zeta(s)$ and hence $\Gamma[A]$, is unchanged.

The price paid for preserving gauge invariance is (as usual) an intrinsic parity anomaly, i.e., one present even in the limit when the explicitly parity violating fermion mass term is absent. [That parity can be sacrificed for gauge invariance was effectively noted in [12].] Under P, $\lambda_n \longrightarrow -\lambda_n^*$ so that $\zeta^P(s) \neq \zeta(s)$. It is easy to express the parity violating part $\Gamma^{(\mathrm{PV})}[A] = 1/2[\zeta'(0) - \zeta'^P(0)]$ explicitly in terms of the eta function in this limit ($m = 0$). Here

$$\begin{aligned}\zeta(s) - \zeta^P(s) &= (1 - e^{-i\pi s})\left(\sum_{\lambda_n>0}(\lambda_n)^{-s} - \sum_{\lambda_n<0}(-\lambda_n)^{-s}\right) \\ &\equiv (1 - e^{-i\pi s})\eta(s), \end{aligned} \tag{4}$$

so that $\Gamma^{(\mathrm{PV})}[A] = i\pi/2\eta(0)$. At $m = 0$, the continuous part of $\eta(0)$ is given in closed form by the CS action [11,13]; being local means it can be removed by a different choice of regularization. For $m \neq 0$ an expansion in powers of the mass can be presented

$$\begin{aligned}\Gamma^{(\mathrm{PV})}(A) &= \frac{1}{2}\frac{d}{ds}[\zeta(s) - \zeta^P(s)]\bigg|_{s=0} \\ &= i\frac{\pi}{2}\eta(0) - i\sum_{k=0}^{\infty}(-1)^k\frac{m^{(2k+1)}}{2k+1}\eta(2k+1), \end{aligned} \tag{5}$$

while the analogous expansion for the parity-conserving part involves even powers of the mass. Several remarks about (5) are in order. (a) The presence of the odd powers can be understood as a consequence of the behavior of the mass term under parity. Instead, the anomalous contribution $\eta(0)$ (proportional to the even m^0 power) originates in a compensation between vanishing and divergent terms. Similarly, for the parity-preserving part there are, besides the even powers, two other possible contributions in three dimensions, one proportional to m and one to m^3, coming from an analogous compensation.

(b) In explicit computations, the expansion, like its analog for the parity preserving part, must be treated carefully, because, even though gauge-invariant order by order, the coefficients of such expansions are not continuous functionals of the gauge field. [Recall, for example, that $\eta(0)$ jumps by ± 2 when an eigenvalue crosses zero or see the $\mathrm{Im}\Gamma[A]$ form in the example below.] The total effective action is, instead, a continuous functional. (c) It would be interesting to compare our mass expansions with the one presented in [13], obtained from low and high temperature limits in four dimensional gauge theories.

For concrete illustrations of how the perturbative non-invariance paradox is circumvented, let us now consider some explicit examples of actions and large gauge transformations both in the Abelian and non-Abelian sectors. The simplest is the pure S^1 $(0+1)$-dimensional toy model of [14], with Dirac operator $[i(d/dt) + A(t) + im]$ and large transformations obeying $f(\beta) - f(0) = 2\pi n$. Charge conjugation $A \longrightarrow -A$ plays the role of parity, which is violated by m, all as in $(2+1)$. Both the eigenvalues and $\zeta(s)$ can be obtained exactly in terms of the average $a = (1/2\pi)\int_0^\beta A(t)\,dt$. We give only the final result here, for N charged fermions:

$$\begin{aligned}\exp[-\Gamma(A)] &= \left\{2\left[\cosh\left(\frac{\beta m}{2}\right)\cos\pi a - i\sinh\left(\frac{\beta m}{2}\right)\sin\pi a\right]\exp\left(i\pi a - \frac{\beta m}{2}\right)\right\}^N \\ &\equiv [\exp(-\beta m + 2\pi i a) + 1]^N.\end{aligned} \tag{6}$$

Note that with our regularization, the action depends on a only via the S^1 holonomy $\exp(2\pi i a)$. Expanding (6) in terms of $\sin k\pi a$ and $\cos k\pi a$ shows the consistency of this result with the general expression (1). A large transformation $a \longrightarrow a + 1$ leaves (6) invariant for any N, even or odd, through a sign cancellation between the separate factors in the middle term. Note the necessary presence of an "intrinsic" charge conjugation anomaly even at $m = 0$: $\mathrm{Im}\Gamma[A] = iN(a - [a])$. This is what allows us to preserve large gauge invariance independently of N. Had we opted instead (as in [14]) for the $(0+1)$ equivalent of the more usual, parity preserving (here C preserving), regularization the $\exp(iN\pi a)$ factor would have been missing and only even N would have kept invariance. The non-Abelian $(0+1)$ scheme is not instructive, essentially because there is no equivalent of the Abelian CS, $\int A$.

A more realistic $(2+1)$ example is the U(1) field

$$A_\mu(t,\mathbf{r}) \equiv \left(\frac{2\pi}{\beta}\,a, \mathbf{A}(\mathbf{r})\right), \tag{7}$$

where a is a flat connection along S^1. $\mathbf{A}$ lives on Σ, with nonvanishing, necessarily integer, flux $\Phi(F) = n$. We concentrate on large transformations $a \longrightarrow a + 1$, although in higher genus Σ one could also have large transformations affecting $\mathbf{A}$. Because of the time independence, we have a tractable eigenvalue equation for λ_n. After some work, it follows that the effective action factorizes into two $(0+1)$-dimensional contributions like (6) and a reduced expression depending on $\mathbf{A}$, Σ, and the holonomy $\exp(2\pi i a)$,

$$\begin{aligned}\exp[-\Gamma(A)] &= [\exp(-\beta m + 2\pi i a) + 1]^{\nu_+}[\exp(-\beta m - 2\pi i a) + 1]^{\nu_-} \\ &\times \left|\prod_{\mu_k}[1 + \exp(-\beta\sqrt{\mu_k^2 + m^2} + 2\pi i a)]\right|^2 \exp[2\pi\zeta_{(\beta^2/4\pi^2)(\hat{\not{D}}^2+m^2)}(-1/2) - (\nu_+ + \nu_-)m\beta].\end{aligned} \tag{8}$$

Here $\hat{\not{D}}$ is the reduced Dirac operator on Σ, μ_k its nonvanishing eigenvalues. [A simple field configuration for which even the μ_k can be computed explicitly is the instanton on the flat unit torus: $A_i = -\pi n\epsilon_{ij}x^j$. Here $\mu_k^2 = 4\pi|nk|$ with degeneracy $2n$, while $2\pi\zeta_{(\beta^2/4\pi^2)(\hat{\not{D}}+m^2)}(-1/2) = n(4\pi n)^{1/2} \times \beta\zeta_H(-1/2, m^2/2\pi n) - (\nu_+ + \nu_-)m\beta$; ζ_H is the Hurwitz function.] The number of positive/negative chiral zero modes $v_\pm$ of $\hat{\not{D}}$ is represented by $\nu_\pm$, with the conventions $(\gamma_5 \mp 1)v_\pm = 0$, and the (parity odd) flux is just $\nu_- - \nu_+$. **(**In $(0+1)$ dimensions, there is no chirality, but an "opposite sign" holonomy can be artificially introduced by considering also fermions subject to a "conjugate" Dirac operator $[-id/dt - A(t) + im]$ which would change the sign of $2\pi i a$ in the last equality of (6).**)** That the infinite product in (8) is convergent follows from the fact that $\mu_k \simeq c\sqrt{|k|}$ [8]. The invariance of (8) under $a \longrightarrow a + 1$ is manifest and its structure is consistent with (1). It is clear that a perturbative (i.e., in powers of a) expansion of (8) loses periodicity in a and hence does not see large invariance order by order. For example, the Chern-Simons term ($I_{\mathrm{CS}} = \pi a n$) has a coefficient $1 - \tanh(\beta m/2)$. The usually quoted coefficient omits the 1 that represents the intrinsic parity-anomaly price of our gauge-invariant regularization and hence persists at $m = 0$. There is actually an ambiguity in its sign [reflecting the choice of cut in (3)], also present in other regularizations—for example, through the factor $\lim_{M\longrightarrow\pm\infty}\mathrm{sign}(M)$ in Pauli-Villars. Irrespective of a expansion, the large m limit of Γ is delicate: with our intrinsic anomaly choice (gauge preserving), we find $\Gamma(A) - \Gamma(0) \longrightarrow (2,0)I_{\mathrm{CS}}$ as $m \longrightarrow (-,+)\infty$; the parity-reversing choice of cut in (3) would yield $(0, -2)$. Any other choice of intrinsic ($m = 0$) anomaly coefficient would, of course, translate these limiting values. These asymptotic properties are independent of the background.

The analogous finite temperature "problem" arises in the context of the non-Abelian theory as well. At zero temperature the loop correction preserves the integer nature of the Chern-Simons coefficient [5], but at finite temperature a puzzling temperature dependence appears [6]. However, the general discussion presented above can be shown to extend naturally to the non-Abelian case, assuring the gauge invariance of the action. To illustrate this, consider the simplest, formally non-Abelian, generalization of the U(1) instanton field considered above: a constant magnetic SU(2) field $F_{ij}^b = 2\pi n \epsilon_{ij} f^b$ on $S^1 \times T^2$, whose gauge potential is $A_\mu^b \equiv [(2\pi/\beta)a, -\pi n \epsilon_{ij} x^j] f^b$, where f^b is a unit color vector and n an integer. Despite appearances, the relevant mechanism here is actually quite different from the Abelian case. There the spectral asymmetry entailing the parity anomaly was governed by the flux $\Phi(F)$ on Σ: geometrically, $\Phi(F)$ represents a nonvanishing Chern class for the reduced two-dimensional field. But the Chern class of a $D = 2$ non-Abelian gauge field vanishes: the asymmetry of the spectrum is not due to the difference in chirality of the zero modes of the reduced Dirac operator on T^2 (the kernel being chirally symmetric), but rather to their different structure as multiplets of SU(2). Consequently, the determinant yields the Abelian result, with $\nu_\pm$ replaced by $2\nu_\pm$. To see this, imagine aligning f^b along, say, the three direction. Then the eigenvalue problem effectively splits into two U(1)'s coupled, respectively, to $\pm A$, so that we just get a doubling of the one-component Abelian result. [For SU(N), one would align f^b along the Cartan sub-algebra, thereby again splitting into various Abelian sectors, with different charges, in a well-defined way.] In this non-Abelian context, the general characteristics we have considered here such as parity anomalies and large gauge-invariance persist at zero temperature and have been discussed, with explicit examples in [15]

In conclusion, we have shown that the apparent large gauge anomalies resulting from a perturbative expansion of the full effective action are due to the more complicated (order-violating) nature of the Ward identities when a nontrivial homotopy is present, the action itself being fully gauge invariant with suitable regularization, one that necessarily entails parity anomalies. This has been illustrated by explicit Abelian and non-Abelian field configurations. Details will be given elsewhere.

This work is supported by NSF Grant No. PHY-9315811, in part by funds provided by the U.S. DOE under cooperative agreement No. DE-FC02-94ER40818, and by INFN, Frascati, Italy.

[1] J. D. Lykken, J. Sonnenschein, and N. Weiss, Int. J. Mod. Phys. A **6**, 5155 (1991).

[2] S. Deser, R. Jackiw, and S. Templeton, Phys. Rev. Lett. **48**, 975 (1982); Ann. Phys. (N.Y.) **140**, 372 (1982).

[3] S. Elitzur, G. Moore, A. Schwimmer, and N. Seiberg, Nucl. Phys. **B326**, 108 (1989); V. P. Nair and M. Bos, Phys. Lett. B **223**, **61** (1989); G. Dunne, R. Jackiw, and C. Trugenberger, Ann. Phys. (N.Y.) **194**, 197 (1989).

[4] A. Niemi and G. W. Semenoff, Phys. Rev. Lett. **51**, 2077 (1983).

[5] R. Pisarski and S. Rao, Phys. Rev. D **32**, 2081 (1985).

[6] R. Pisarski, Phys. Rev. D **35**, 664 (1987).

[7] K. Babu, A. Das, and P. Panigrahi, Phys. Rev. D **36**, 3725 (1987); I. Aitchison and J. Zuk, Ann. Phys. (N.Y.) **242**, 77 (1995); N. Bralić, C. Fosco, and F. Schaposnik, Phys. Lett. B **383**, 199 (1996); D. Cabra, E. Fradkin, G. Rossini, and F. Schaposnik, Phys. Lett. B **383**, 434 (1996).

[8] R. T. Seeley, Proc. Sympos. Pure Math. **10**, 288 (1968); P. B. Gilkey, *Invariance Theory, the Heat-Equation and the Atiyah-Singer Index Theorem* (CRC Press, Boca Raton, 1995).

[9] S. W. Hawking, Commun. Math. Phys. **55**, 133 (1975).

[10] R. E. Gamboa Saravi, G. L. Rossini, and F. A. Schaposnik, Int. J. Mod. Phys. **A11**, 2643 (1996).

[11] L. Alvarez-Gaumé, S. Della Pietra, and G. Moore, Ann. Phys. (N.Y.) **163**, 288 (1985); A. Niemi, Phys. Rev. Lett. **57**, 1102 (1986).

[12] A. N. Redlich, Phys. Rev. Lett. **52**, 18 (1984); Phys. Rev. D **29**, 2366 (1984).

[13] A. R. Rutheford, Phys. Lett. B **172**, 187 (1986).

[14] G. Dunne, K. Lee, and C. Lu, Phys. Rev. Lett. **78**, 3434 (1997).

[15] S. Forte, Nucl. Phys. B **288**, 252 (1987).